

ORIGINAL ARTICLE

Mitochondrial genomes of four pierid butterfly species (Lepidoptera: Pieridae) with assessments about Pieridae phylogeny upon multiple mitogenomic datasets

Lan Nie¹, Yunliang Wang¹, Dunyuan Huang², Ruisong Tao¹, Chengyong Su¹, Jiasheng Hao¹*, Chaodong Zhu³*

¹Laboratory of Molecular Evolution and Biodiversity, College of Life Sciences, Anhui Normal University, Wuhu 241000, China

²College of Life Sciences, Chongqing Normal University, Chongqing 400047, China

³Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

*Corresponding authors, E-mails: jshaonigpas@sina.com; zhucd@ioz.ac.cn

Abstract In this study, the complete mitochondrial genomes of four pierid butterfly species, namely *Battlia butleri*, *Talbotia naganum*, *Pontia callidice* and *P. daplidice*, were newly sequenced and characterized. Meanwhile the phylogenetic relationships of the main Pieridae lineages covering 22 pierid butterfly species, were reconstructed with Bayesian inference and maximum likelihood methods based on different mitogenomic datasets, including the concatenated 13 PCGs, 13 PCGs + 2 rRNAs, 2 rRNAs, 2rRNAs + 22 tRNAs and 22 tRNAs sequences, respectively. Our results of mitogenomic analysis showed that the four mitogenomes were 15,124, 15,155, 15,109, 15,124 bp in size, with the gene orders and arrangements identical to all other butterflies determined. Our results of phylogenetic analyses upon protein coding genes, rRNA genes or their sequence combinations indicated that the three pierid subfamilies of this study were all monophyletic, with their relationships of being (Dismorphiinae, (Pierinae, Coliadinae)); the phylogenetic relationship of Pierini of this study was (*Pieris*, (*Baltia*, (*Talbotia*, *Pontia*))) + (*Prioneris*, (*Delias*, *Aporia*)). Additionally, our analyses suggested that only mitogenomic tRNA sequence datasets were not ready to be utilized in resolving the deeper phylogeny, whereas somewhat suitable to be applied in clarifying the phylogenies of closely related species, of the pierid butterflies.

Keywords Mitogenomes, phylogeny, Pieridae.

1 Introduction

The family Pieridae is one of the largest butterfly groups which is traditionally divided into four subfamilies (Pseudopontiinae, Dismorphiinae, Coliadinae, Pierinae) covering about 85 genera, 1200 living species distributed nearly all around the world (Singh *et al.*, 1998; Braby *et al.*, 2006; Wahlberg *et al.*, 2014; Ding & Zhang, 2016). The subfamily Pseudopontiinae contains only one genus *Pseudopontia* with probably five species distributed in tropical Africa; the subfamilies Dismorphiinae contains seven genera covering about sixty species mostly distributed in only the Neotropical areas; the subfamily Coliadinae contains about eighteen genera covering about 220 species, while the subfamily Pierinae contains nearly sixty genera and exceeds 900 species, and this two subfamilies are distributed worldwide (Braby & Trueman, 2006; Cao *et al.*, 2016).

Up to the present, the phylogeny of Pieridae, especially those of a few higher and many lower taxa, are not fully resolved

(Ehrlich, 1958; Ehrlich & Ehrlich, 1967; Pollock *et al.*, 1998; Braby *et al.*, 2006; Wahlberg *et al.*, 2014). For example, some previous studies suggested that the Pseudopontiinae and Dismorphiinae were reciprocally monophyletic sister groups (Ehrlich, 1958; Braby *et al.*, 2006), whereas others revealed that the Pseudopontiinae should be considered as an intermediate between the Dismorphiinae and Pierinae *sensu lato* (*s.l.*) (including the Coliadinae, formerly the tribe Coliadini) (Clench, 1955; Cao *et al.*, 2016). In addition, the internal phylogeny of the tribe Pierini (the relationship among the genera *Pieris*, *Baltia*, *Pontia* and *Talbotia*) has been far away from a consensus. Braby *et al.* (2006) revealed that the Pierini was a monophyletic group, with the relationship of (*Talbotia*, ((*Baltia*, *Pontia*), *Pieris*)) by using criteria of multiple gene sequence data (*EF-1 α* , *COI*, wingless, 28S rRNA) and morphological characters; Xu *et al.* (2007) obtained the similar results using *COI* and *Cytb* gene sequence data, in which the genus *Baltia* is not involved. However, Ding and Zhang (2016) showed that the phylogenetic relationship among the four genera of the Pierini family was ((*Baltia*, *Pontia*), (*Talbotia*, *Pieris*)) by using combined mitochondrial *COII*, *NDI*, *Cytb*, 16S rRNA and nuclear 28S rRNA sequence data.

Insect mitochondrial genome (mitogenome) is usually double-stranded, circular molecule, ranging in size from 14 to 20 kb (Brown & Wilson, 1979). It usually contains a conserved set of 37 genes, including three cytochrome oxidase (*COI-III*), seven *NADH dehydrogenase* (*ND1-6*, *ND4L*), two *ATPase* (*ATP6*, *ATP8*), one *cytochrome b* (*Cytb*), two rRNA (*12S rRNA* and *16S rRNA*), 22 tRNA genes, and a noncoding AT-rich region (Shadel & Clayton, 1993; Boore, 1999; Jiang *et al.*, 2009; Cameron, 2014). Due to its relatively small size, simple organization, high evolutionary rate and unambiguous orthology, mitogenome has become one of the most powerful tools and was widely used in the phylogenetic analysis and other relevant research areas (Shadel & Clayton, 1993; Boore, 1999; Dowton *et al.*, 2002; Liu *et al.*, 2013; Sun *et al.*, 2016).

In this study, we newly determined the complete mitochondrial genomes of four pierid butterfly species, namely *Baltia butleri*, *Talbotia naganum*, *Pontia callidice* and *P. daplidice*. We attempted to conduct the phylogenetic analysis of the Pieridae by combining with other available mitogenomes of 18 pierid butterflies with different analytic methods based on multiple mitogenomic concatenated dataset (13 PCGs, 13 PCGs+2 rRNAs, 22 tRNAs, 2 rRNAs, 22 tRNAs+2 rRNAs), respectively, in order to clarify the phylogenetic relationships of the Pieridae lineages in detail on different levels. Additionally, we attempted to evaluate the usage feasibility of mitogenomic RNA sequences, including its rRNAs and tRNAs, in the butterfly phylogenetic analysis at different taxonomic levels.

2 Materials and methods

2.1 Sample collection

Adult individuals of *Talbootia naganum* were collected from the Jiulianshan Mountains, Jiangxi Province, China, in August 2009. *Baltia butleri*, *Pontia callidice*, and *P. daplidice* were collected from the Qilianshan Mountains, Qinghai Province, China, in July 2016. After samples were collected and identified by JS Hao, the fresh leg tissues were preserved in 100% ethanol for fixation and stored at -20°C until follow-up experiments.

2.2 DNA extraction and PCR amplification

Total genomic DNA was extracted from leg muscle of each adult individual by the glass powder method with minor modifications (Hao *et al.*, 2005). Three partial fragments of *COI*, *ND4*, and *Cytb* were amplified by using standard short primers from Simon *et al.* (1994), other primers for the amplifications of long and short fragments were designed by Primer Premier 5.0 (Singh *et al.*, 1998) through multiple sequence alignments of butterfly mitogenomes using Clustal X1.8 (Thompson *et al.*, 1997). All the primers were synthesized by Sangon Biotechnology Co. Ltd. (Shanghai, China) (Table 1). Long PCR amplification reactions were performed under the following conditions: an initial denaturation at 95°C for 5 min; 30 cycles of denaturation (denaturing at 95°C for 50 s, annealing at 50–55°C for 50 s, extension at 68°C for 150 s), and a final extension step at 68°C for 10 min. The PCR products were separated by 1.2% agarose gel electrophoresis, purified by using a 3S Spin PCR Product Purification Kit (Sangon Biotechnology Co. Ltd., Shanghai) and sequenced on an ABI-377 automatic DNA sequencer. For each long PCR product, the complete, double-stranded sequence was determined by the “primer walking” method.

2.3 Sequence analysis

The raw sequences files were edited and assembled with BioEdit v7.0 (Hall, 1999). The concatenated amino acid sequences of the 13 PCGs, rRNA genes and the A+T-rich regions were determined through the alignment of the sequences

with the homologous regions in other Pierinae mitogenome sequences using MEGA7.0 (Kumar *et al.*, 2016). The tRNA genes were identified by using the tRNAscan-SE 1.21 (Lowe & Eddy, 1997), and the putative tRNAs that could not be identified by the tRNAscan-SE 1.21 were retrieved by comparison with other pierid butterflies. The rRNA secondary structures were predicted by tRNAscan-SE 1.21 (Lowe & Eddy, 1997) and modified through comparison with other related butterfly species (Cameron & Whiting, 2008; Sun *et al.*, 2015). The amino acid sequences of the PCGs were translated via the invertebrate mtDNA genetic code. The nucleotide composition, frequency of codon usage and the relative synonymous codon usage (RSCU) of the PCGs were calculated by MEGA7.0 (Kumar *et al.*, 2016), while the tandem repeats in the A+T-rich regions were predicted by the Tandem Repeats Finder (Benson, 1999).

Table 1. List of primer used for the amplification of the four pierid butterfly mitogenome.

Short primer	Primer sequence (F/R) 5'–3'	AT*	Long primer	Primer sequence (F/R) 5'–3'	AT*
<i>COI</i>	GGTCAACAAATCATAAAGATAT TAAACTTCAGGGTGACCAAAAAAT	51.7°C	<i>COI-COII</i>	CCCCTCTGTAACCCTT GCAAACTATGATTGGCTC	49.4°C
<i>COII</i>	GAGACCATTA CTGCTTT CAGTCACT CTAATATGGCAGATTATATGTATTGG	52.8°C	<i>COII-COIII</i>	AACGATTCCCTTCCCTG ATGATGCTGCTGCTTC	49.8°C
<i>Cytb</i>	GTCCTACCATGAGGTCAAATATC TTCAACTGGTCGTGCTCCAATTCA	52.2°C	<i>COIII-ND5</i>	CACCACTTTGGCTTTGA GATTTCTGGGGTTGACTAT	49.1°C
<i>NDI</i>	CGTAAAGTCCTAGGTTATATTCAGATTTCG ATCAAAAGGAGCTCGATTAGTTTC	52°C	<i>ND5-ND4</i>	ATAAACCTTGCCCT GATTCTTCAGTTGCTC	42.8°C
<i>lrRNA</i>	CCGGTTTGAACCTCAGATCACGT CGCCTGTTTATCAAAAACAT	46°C	<i>ND4-Cytb</i>	AATAAACCATACCCACCT TACAGCGAATCCTCCTC	49.3°C
<i>COIII</i>	GTTCTGAGATTT CAGGTAA TTACTAATAAATCATTTC	46.5°C	<i>Cytb-srRNA</i>	ATTCCTCCATCAAACAG AAAGTCTAATCTTCCCAC	45.1°C
<i>ND4</i>	ATATTTTGTAYHCCACAAATC CAGGTTCAATAATTTTAGC	46.6°C	<i>srRNA-ND2</i>	GAAACACTTTCCAGTACCT CTAAACCAATTCAACATCC	49.6°C

*Annealing temperature

2.4 Phylogenetic analysis

The ingroup taxa of this study for the phylogenetic analyses included a total of 22 available mitochondrial genome data representing three subfamilies (four newly determined, and 18 directly extracted from GenBank). Two Nymphalidae butterfly species, *Fabriciana nerippe* and *Issoria lathonia* were used as outgroups (Table 2). The phylogenetic trees were reconstructed with the Bayesian inference (BI) and maximum likelihood (ML) methods based on the concatenated 13 PCGs, 13 PCGs + 2 rRNAs, 2 rRNAs, 2 rRNAs + 22 tRNAs and 22 tRNAs sequence datasets, respectively. BI analysis was conducted in MrBayes 3.2 (Ronquist *et al.*, 2012) with the best-fit partitioning scheme and partition-specific substitution model selected using the 'greed' algorithm in the program PartitionFinder v1.1.1 (Lanfear *et al.*, 2012) under the Bayesian information criterion (BIC) (Table 3). In the analysis, two parallel analyses with four independent Markov chains were run for about 1,000,000 generations by sampling every 100 generations, and chain convergence was considered to be reached when the average standard deviation of split frequencies fell below 0.01; after discarding the first 25% burn-in trees, the 50% majority rule consensus tree was reconstructed by summarizing the remaining trees with posterior probability (PP) values on each tree node. ML analysis was conducted by using the raxmlGUI.5b1 interface (Silvestro & Michalak, 2012) of RAxML v7.2.6 (Stamatakis, 2006) with the evaluated best-fit model of GTRGAMMAI; the bootstrap support (BS) nodal values of each tree were obtained via rapid bootstrapping with 1,000 pseudoreplicates. In both BI and ML analyses, the data partitions of different combined datasets were set as unlinked, with the prior rate of BI set as variable (Ding & Zhang, 2016; Su *et al.*, 2017).

3 Results

3.1 General mitogenomic features

Table 2. Summarized mitogenomic characteristics of the 24 butterfly species investigated in this study.

Species	Whole genome		PCG		lrRNA		srRNA		A+T-rich region		Accession number*	Reference
	Size(bp)	A+T(%)	No.codons	A+T(%)	Size (bp)	A+T(%)	Size(bp)	A+T(%)	Size (bp)	A+T(%)		
<i>Pieris melete</i>	15,140	79.8	3,714	78.5	1,319	83.4	777	86.9	351	88.0	NC_010568	(Yang <i>et al.</i> , 2009)
<i>Gonepteryx mahaguru</i>	15,221	80.9	3,712	79.4	1,334	84.5	778	85.7	385	95.1	NC_026837	(Yang <i>et al.</i> , 2014)
<i>Pieris canidia</i>	15,153	79.7	3,712	78.2	1,323	83.9	767	84.2	418	91.2	NC_026532	(Fang <i>et al.</i> , 2015)
<i>Hebomoia glaucippe</i>	15,701	79.9	3,715	78.0	1,323	83.7	777	85.2	899	92.2	NC_021123	(Hao <i>et al.</i> , 2013)
<i>Pieris rapae</i>	15157	79.8	3,720	78.2	1,330	84.0	764	85.0	393	91.6	NC_015895	(Mao <i>et al.</i> , 2010)
<i>Delias hyparete</i>	15,186	79.8	3,703	78.4	1,311	80.7	786	78.2	377	92.0	NC_020428	(Shi <i>et al.</i> , 2012)
<i>Aporia crataegi</i>	15,140	81.3	3,708	79.9	1,326	85.4	779	85.5	354	95.2	NC_018346	(Park <i>et al.</i> , 2012)
<i>Catopsilia pomona</i>	15,142	81.3	3,724	80.0	1,332	85.2	779	85.1	313	97.1	NC_022687	(Hao <i>et al.</i> , 2014)
<i>Leptidea morsei</i>	15,122	80.2	3,713	79.2	1,337	84.3	764	83.2	356	89.3	NC_022686	(Hao <i>et al.</i> , 2014)
<i>Gonepteryx rhamni</i>	15,203	80.3	3,714	78.8	1,329	84.4	779	84.9	371	94.9	NC_026046	Unpublished
<i>Aporia intercostata</i>	15,144	80.4	3,717	79.0	1,322	84.9	771	82.7	368	95.9	NC_025273	Unpublished
<i>Anthocharis bambusarum</i>	15,180	80.2	3,721	78.9	1,348	84.0	780	84.1	318	90.3	NC_025274	Unpublished
<i>Colias erate</i>	15,184	81.3	3,719	80.0	1327	84.7	767	85.2	364	95.1	NC_027253	(Wu <i>et al.</i> , 2015)
<i>Aporia hippia</i>	15,154	79.5	3,715	77.9	1,326	84.2	771	84.6	382	94.7	NC_033889	(Cao <i>et al.</i> , 2016)
<i>Mesapia peloria</i>	15,159	81.1	3,717	79.7	1,323	85.0	772	85.2	378	94.7	NC_033891	(Cao <i>et al.</i> , 2016)
<i>Prioneris clemathe</i>	15,115	80.7	3,713	79.3	1,342	84.9	752	84.6	341	95.3	not uploaded	Unpublished
<i>Aporia martineti</i>	15,192	80.6	3,717	79.1	1,326	84.6	774	84.8	418	93.1	NC_033890	(Cao <i>et al.</i> , 2016)
<i>Aporia bieti</i>	15,147	80.0	3,718	78.6	1,328	84.4	774	85.2	362	92.8	NC_033888	(Cao <i>et al.</i> , 2016)
<i>Baltia butleri</i>	15,124	79.7	3,715	78.3	1,322	83.6	774	84.6	361	91.4	MH380204	this study
<i>Talbotia naganum</i>	15,155	79.8	3,717	78.3	1,321	83.6	773	85.5	387	95.0	MH380205	this study
<i>Pontia callidice</i>	15,109	79.8	3,706	78.5	1,312	83.9	774	84.6	362	90.9	MH380206	this study
<i>Pontia daplidice</i>	15,124	79.9	3,706	78.4	1,319	84.0	774	85.2	374	92.7	MH380207	this study
<i>Issoria lathonia</i>	15,171	81.2	3,719	79.8	1,319	84.3	771	85.1	336	95.8	NC_018030	Unpublished
<i>Fabriciana nerippe</i>	15,140	80.9	3,718	79.7	1,321	84.5	773	84.9	329	95.5	NC_016419	(Kim <i>et al.</i> , 2011)

*The data of *Prioneris clemathe* is from the laboratory data, others were download from GenBank.

The four complete mitogenomes *Baltia butleri*, *Talbotia naganum*, *Pontia callidice*, and *P. daplidice* all contained 13 PCGs (*COI*–*COIII*, *ATP6*, *ATP8*, *ND1*–*6*, *ND4L*, *Cytb*), 22 putative tRNA genes, 2 rRNA genes (*lrRNA* and *srRNA*), and a non-coding control area called A+T-rich region, with the sizes of 15,124, 15,155, 15,109, and 15,124 bp, respectively (Fig. 1). The gene order and arrangement of the four mitogenomes are also similar to those of other butterflies. Four PCGs (*ND5*, *ND4*, *ND4L*, *ND1*), 8 tRNAs (*trnQ*, *trnC*, *trnY*, *trnF*, *trnH*, *trnP*, *trnL*, *trnV*), and 2 rRNAs (*lrRNA* and *srRNA*) are coded on the minority-strand, whereas the remaining 23 genes coded on the majority-strand. In addition, like other butterfly mitogenomes, the three tRNAs between the A+T-rich region and *ND2*, are arranged in the pattern of *tRNA^{Met}/tRNA^{Ile}/tRNA^{Gln}*, different from that found in the ancestral insects (*tRNA^{Ile}/tRNA^{Gln}/tRNA^{Met}*) (Fig. 1).

The nucleotide compositions of the four mitogenomes are significantly A+T biased, with the values of 79.7, 79.8, 79.8 and 79.9, respectively (Table 2). The cases are similar also for each mitogenomic gene or noncoding region of the four mitogenomes. For example, the A+T-rich regions in turn are 91.4%, 95.0%, 90.9%, and 92.7%, while the PCGs (exclusive

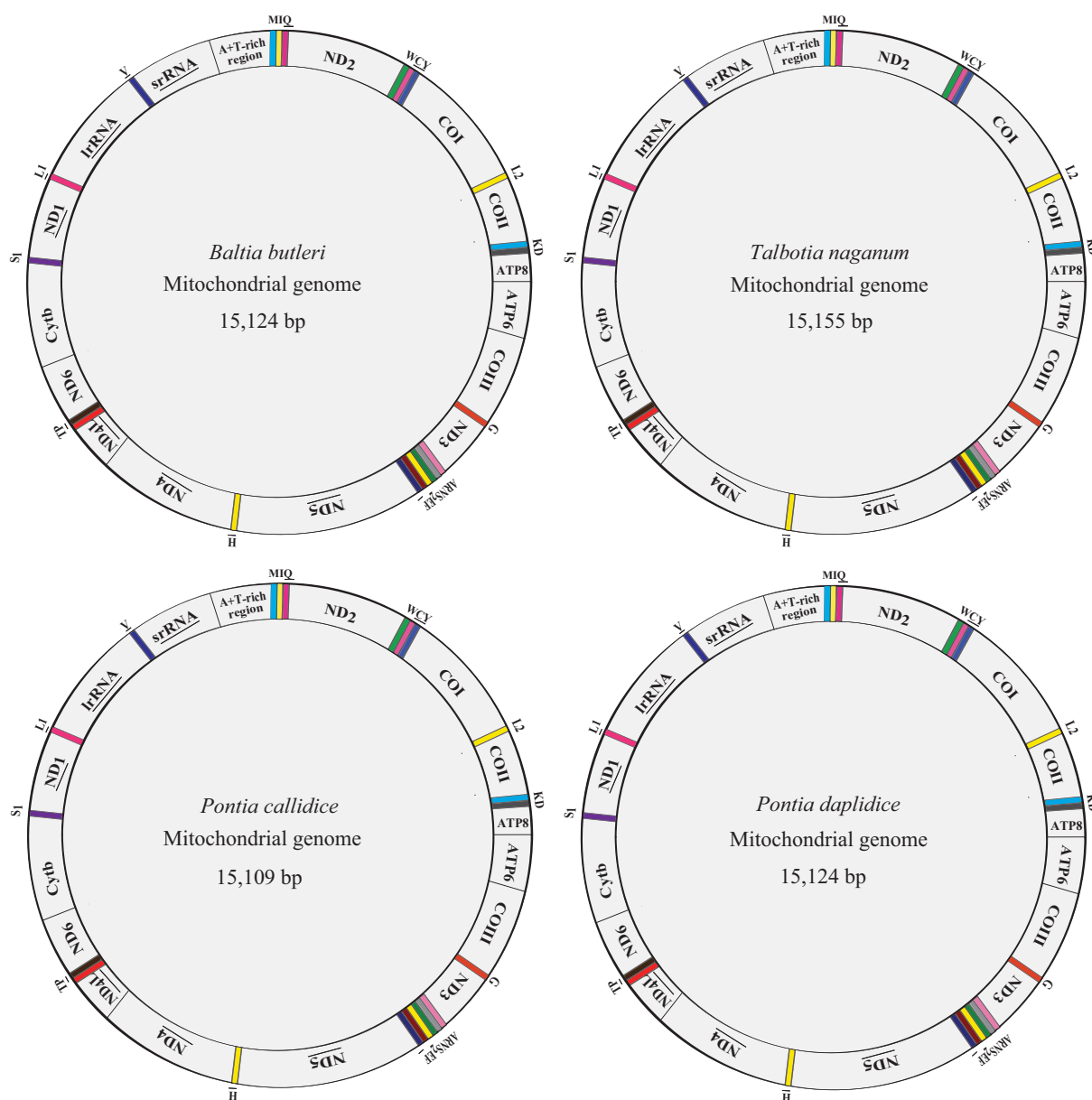


Figure 1. Circular map of *Baltia butleri*, *Talbotia naganum*, *Pontia callidice*, *Pontia daplidice* mitochondrial genome. *COI*, *COII*, and *COIII* refer to the cytochrome oxidase subunits; *Cytb* refers to cytochrome B; *ATP6* and *ATP8* refer to subunits 6 and 8 of F0 ATPase; *ND1*–*6* refers to the components of NADH dehydrogenase. The tRNAs locations are marked by the color blocks and labeled by the IUPAC-IUB single letter amino acid code. L1, L2, S1, and S2 denote *tRNA^{Leu}* (CUN), *tRNA^{Leu}* (UUR), *tRNA^{Ser}* (AGN), and *tRNA^{Ser}* (UCN), respectively. The non-underlined genes are transcribed on the majority strand whereas the underlined genes are transcribed on the minority strand.

of the stop codon) are 78.3%, 78.3%, 78.5%, and 78.4% in A+T content, respectively (Table 2).

3.2 PCGs

The PCGs of the four species (*Baltia butleri*, *Talbotia naganum*, *Pontia callidice* and *P. daplidice*) harbor a total of 3715, 3717, 3706, and 3706 codons, respectively (Table 2). Twelve of the 13 PCGs start with the typical ATN codons, whereas the *COI* begin with CGA (Table 4). This CGA codon of *COI* is highly conserved throughout all the lepidopteran

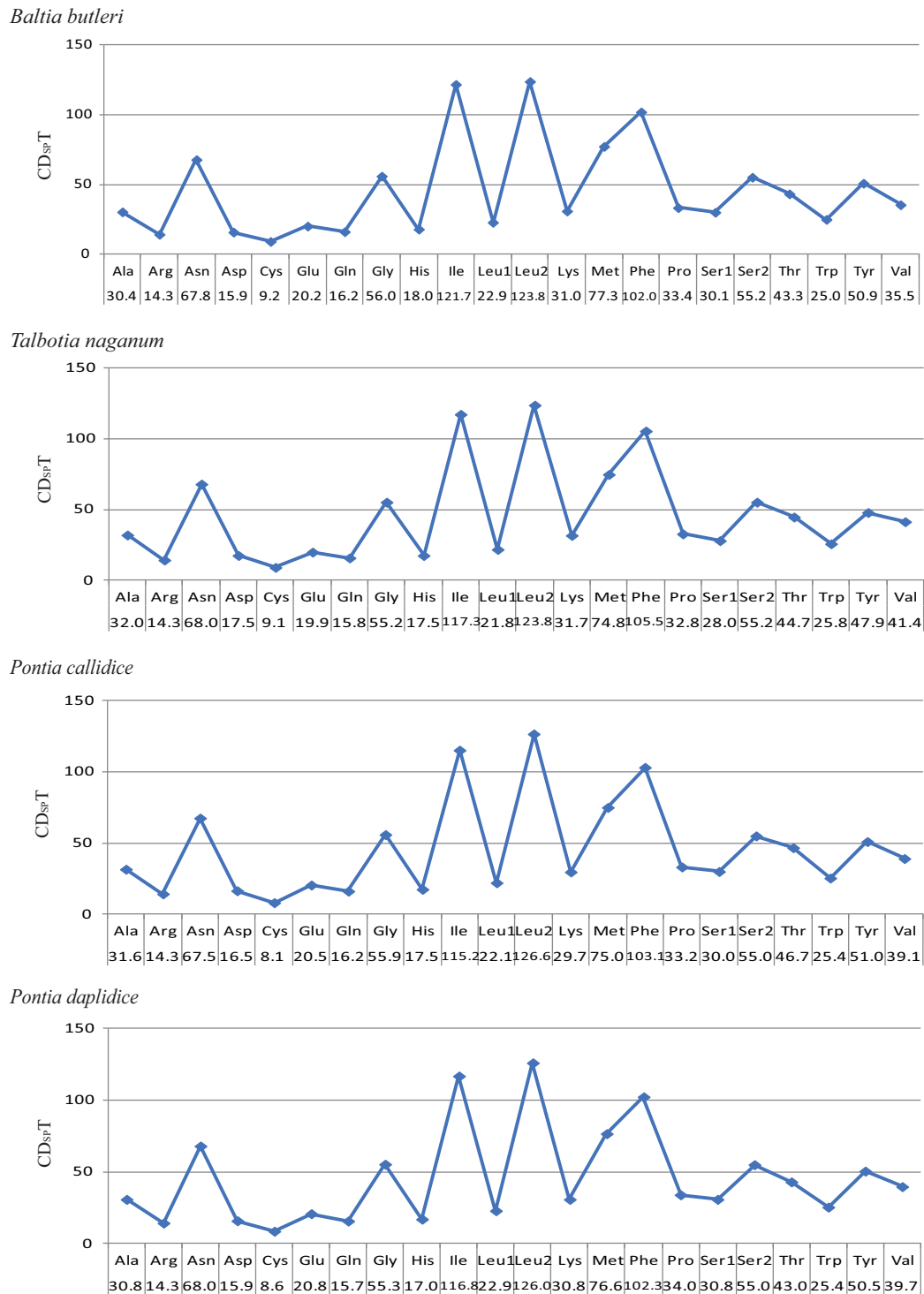


Figure 2. Codon distribution in the mitogenomic PCGs of the four pierid butterfly species newly determined in this study. The Y-axis represents the total number of a codon (CDsp T, codons per thousand codons) The X-axis represents codon families.

groups acting as a synapomorphic character of the Lepidoptera (Kim *et al.*, 2009). For example, in *Baltia butleri* and *Talbotia naganum*, the *ATP8* starts with ATT, whereas in *Pontia callidice* and *P. daplidice*, the *ATP8* starts with the codon ATC. The majority of PCGs of the stop with standard codon TAN, while a few start with incomplete codons, for example, the *ND4* ends with TA whereas the *COI* and *COII* end with T (Table 4).

The A+T contents of the 13 PCGs of the four species are 78.3%, 78.3%, 78.5%, and 78.4%, respectively, fallen with the range of other known pierid butterflies from 78% in *Hebomoia glaucippe* to 80% in *Catopsilia pomona* (Table 2). In addition, for A+T content of the PCG three codon sites, the third codon position is considerably higher than the first and the



Figure 3. Relative Synonymous Codon Usage (RSCU) of the four pierid butterfly mitogenomes newly determined in this study.

second. The frequency of codon usage and the relative synonymous codon usage (RSCU) analyses showed that TTA (Leu), ATT (Ile), TTT (Phe), ATA (Met) and AAT (Asn) are the five most frequently used codons, indicating a bias for the protein-coding gene AT; the NNU and NNA codon RSCU for each of these sites is greater than 1, which indicates that the third codon often leads to the transcription of U and A. The codon bias is positively correlated with the A+T content of the third codon of the protein-coding gene. These results suggest that codon use is strongly biased in all species. All used codons were present in the PCGs of the *Battlia butleri* mitogenome, except for CGC and AGG. This is similar to codon usage in *Talbotia naganum*, *Pontia callidice* and *P. daplidice*, which lack CGC, AGG, AGC and ACG; CGC, CCG and ACG; CGC, CAG, CCG, AGG and UGG, respectively (Sun *et al.*, 2016) (Figs 2–3).

3.3 tRNAs and rRNAs

The mitogenomes of the four pierid species all contain 22 tRNA genes. All the tRNAs could form the typical clover leaf secondary structure except *tRNA^{ser}* (AGN), which lacks the dihydrouridine loop, as universally found in other butterfly mitogenomes (Zhang *et al.*, 2013; Hao *et al.*, 2014; Cao *et al.*, 2016). In addition, all the tRNAs harbor base pair mismatches, such as the G·U, U·G, U·U, A·C, and A·A, and these unusual weak-bonded base mismatches might be corrected by RNA editing (Lavrov, Brown, & Boore, 2000) (Figs 4–5). A total of 33 mismatches (3 A-A, 1 A-C, 1 C-U, 1 G-A and 4 U-U) and 23 G-U wobble pairs are detected in 18 tRNA genes of *Battlia butleri*; 27 mismatches (1 A-A, 1 A-C, 1 C-U, 1 G-A and 3 U-U) and 20 G-U wobble pairs are detected in 16 tRNA genes of *Talbotia naganum*; 25 mismatches (2 A-A, 3 A-C and 1 U-U) and 19 G-U wobble pairs are detected in 18 tRNA genes of *Pontia callidice*; 26 mismatches (2 A-A, 2 A-C and 2 U-U) and 20 G-U wobble pairs are detected in 16 tRNA genes of *P. daplidice* (Sun *et al.*, 2016).

All the four mitogenomes contained two ribosomal RNA genes (*lrRNA* and *srRNA*) located between *tRNA^{Leu}* (CUN) and *tRNA^{Val}*, and between *tRNA^{Val}* and the A+T-rich region, respectively (Fig. 1). The four *lrRNAs* are 1322, 1321, 1312, and 1319 bp in turn, each with 83.6%, 83.6%, 83.9%, and 84.0% AT content, respectively; while the *srRNAs* are 774, 773, 774, and 774 bp with 84.6%, 85.5%, 84.6%, and 85.2% AT content, respectively (Table 2). The predicted topologies of the *lrRNA* and *srRNA* secondary structures are generally in congruence with those previously reported from other lepidopterans (Cameron & Whiting, 2008; Sun *et al.*, 2012; Shi *et al.*, 2015; Cao *et al.*, 2016). For example, the four pierid species *lrRNA* contains three domains (labeled I, II and III) with 50 helices, respectively; *srRNA* contains six domains (labeled I, II, III, IV, V, and VI), respectively (Figs 6–13).

3.4 The A+T-rich regions

The A+T-rich regions of the four mitogenomes are 361, 387, 362, and 374 bp, respectively, all located between the *srRNA* and *tRNA^{Met}* (Table 2). The A+T contents of these regions are 91.4%, 95.0%, 90.9%, and 92.7%, respectively, with their values being the highest among all the mitogenomic genes and noncoding sequences (Fig. 1, Table 2). The case of significantly strong A+T bias of this region in pierids ranging from the *Pieris melete* (88.0%) to *Catopsilia pomona* (97.1%) is commonly found in other lepidopterans (Li *et al.*, 2015). Additionally, these regions of the four mitogenomes also harbor some structures characteristic of lepidopterans (Cao *et al.*, 2016), such as a poly-T stretch upstream of the motif ATAGA, the microsatellite-like elements (TA)_n (*n* = 6–10) preceded by the ATTTA motif, *etc.* (Fig. 14).

Table 3. The partition-specific substitution model.

Datasets	Subsets	Best model
13PCGs	<i>ATP6, ATP8, COI, COII, Cytb, ND2, ND3, ND4, ND5, ND6</i> <i>COIII, ND1, ND4L</i>	GTR+I+G GTR+G
13PCGs+2rRNAs	<i>ATP6, ATP8, COI, COII, Cytb, ND2, ND3, ND4, ND5, ND6, 12S, 16S</i> <i>COIII, ND1, ND4L</i>	GTR+I+G GTR+G
2rRNAs	<i>12S, 16S</i>	GTR+I+G
2rRNAs+2tRNAs	<i>12S, 16S, Ile, Lys, Ser2, Thr, Val</i> <i>Ala, Arg, Cys, Met</i> <i>Asn, Gly</i> <i>Asp, Glu, Leu1, Pro, Tyr</i> <i>Gln, His, Phe, Ser1, Trp</i> <i>Leu2</i>	GTR+I+G GTR+G F81+G HKY+I+G HKY+G GTR+I
22tRNAs	<i>Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu1, Leu2, Lys, Met, Phe, Pro, Ser1, Ser2, Thr, Trp, Tyr, Val</i>	GTR+I+G

Table 4. Initiation and termination codons for the 13 protein-coding genes (PCGs) of the 24 butterfly species used in this study. *

Species	Predicted initiation and termination codons												
	<i>ATP6</i>	<i>ATP8</i>	<i>COI</i>	<i>COII</i>	<i>COIII</i>	<i>Cytb</i>	<i>ND1</i>	<i>ND2</i>	<i>ND3</i>	<i>ND4</i>	<i>ND4L</i>	<i>ND5</i>	<i>ND6</i>
<i>Pieris melete</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATA/TAA	ATA/TAA	ATT/TAA	ATT/TAA
<i>Gonepteryx mahaguru</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAA	ATG/TA	ATG/TAA	ATT/T	ATT/TAA
<i>Pieris canidia</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATG/TAA	ATT/TAG	ATT/TAA
<i>Hebomoia glaucippe</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATT/TAA
<i>Pieris rapae</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATG/TAA	ATT/TAG	ATT/TAA
<i>Delias hyparete</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATA/TAA	ATT/T	ATT/TAA
<i>Aporia crataegi</i>	ATG/TAA	ATC/TAA	ATT/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TA	ATG/T	ATT/T	ATC/TAA
<i>Catopsilia pomona</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATC/TAA
<i>Leptidea morsei</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATA/TAA	ATA/T	ATG/TAA	ATT/T	ATT/TAA
<i>Gonepteryx rhamni</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATA/TAA	ATA/TAA	ATG/TAA	ATT/TAA	ATT/TAA	ATA/TA	ATG/TAA	ATT/T	ATT/TAA
<i>Aporia intercostata</i>	ATG/TAA	ATC/TAA	ATG/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/TAA	ATC/TAA
<i>Anthocharis bambusarum</i>	ATG/TAA	ATT/TAA	ATT/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/T	ATT/TAG	ATG/T	ATG/TAA	ATA/TAA	ATC/TAA
<i>Colias erate</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATC/TAA
<i>Aporia hippia</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATA/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATT/TAA
<i>Mesapia peloria</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATC/TAA
<i>Prioneris clemathe</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATA/TAA	ATA/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATT/TAA
<i>Aporia martineti</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATC/TAA
<i>Aporia bieti</i>	ATG/TAA	ATA/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATT/TAA
<i>Baltia butleri</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TA	ATG/TAA	ATT/TAA	ATC/TAA
<i>Talbotia naganum</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TA	ATG/TAA	ATT/TAA	ATT/TAA
<i>Pontia callidice</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TA	ATG/TAA	ATT/TAA	ATC/TAA
<i>Pontia daplidice</i>	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAA	ATC/TAG	ATG/TA	ATG/TAA	ATT/TAG	ATC/TAA
<i>Issoria lathonia</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TA	ATT/TAA	ATT/T	ATG/TA	ATG/TAA	ATT/TAA	ATT/TAA
<i>Fabriciana nerippe</i>	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/T	ATT/TAA	ATG/TA	ATG/TAA	ATT/TA	ATT/TAA

*The data of *Prioneris clemathe* is from the laboratory data of authors, others were download from GenBank.

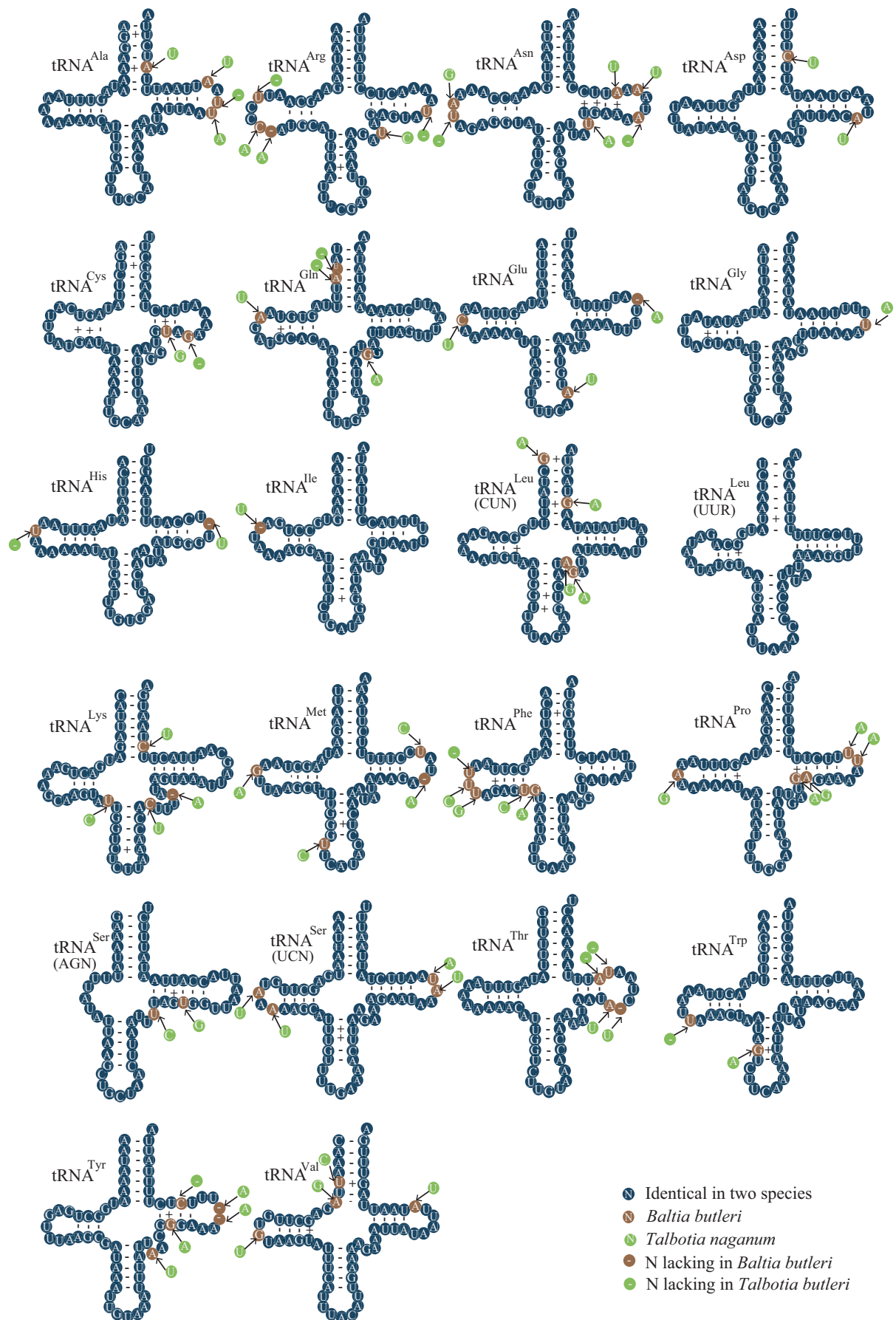
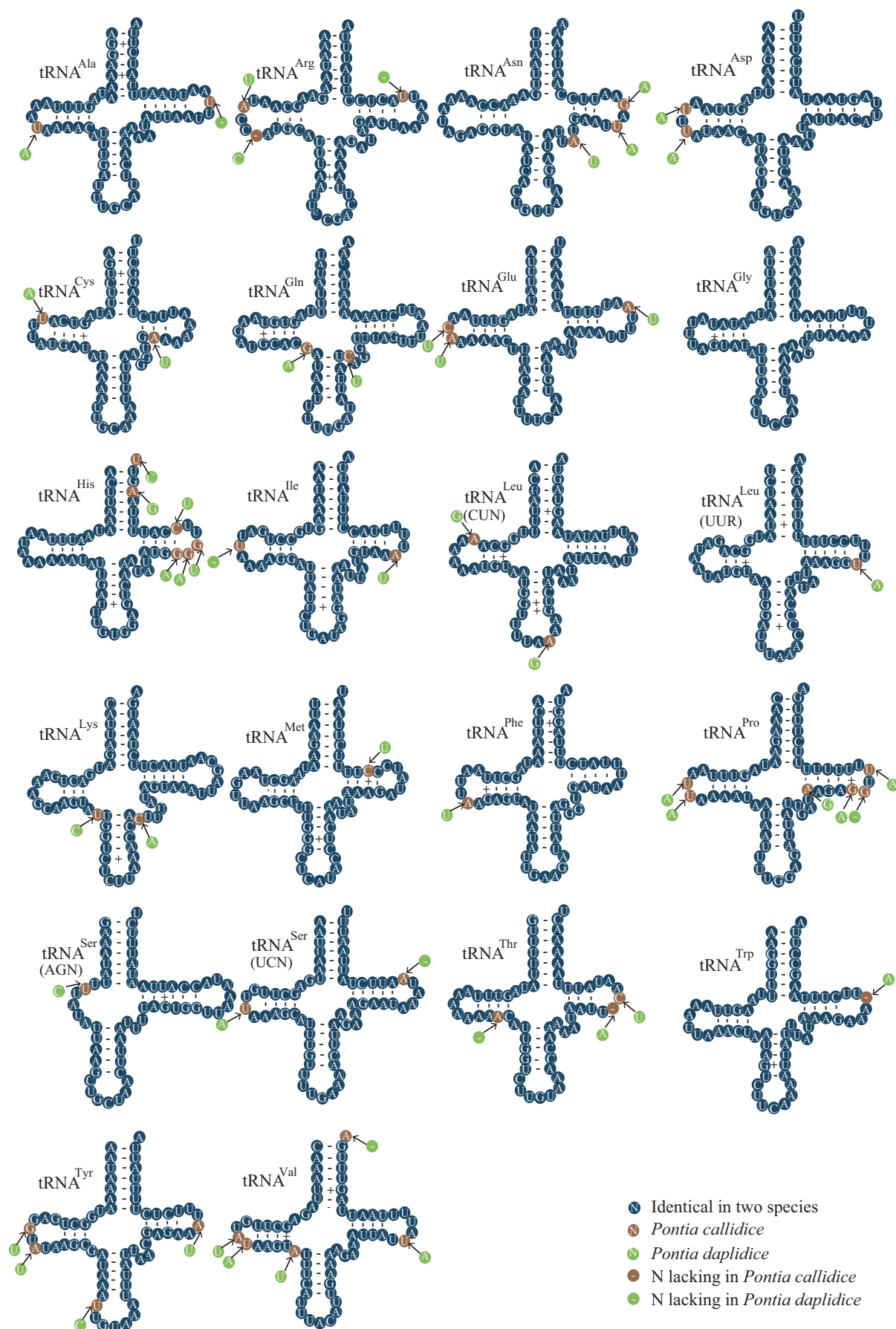


Figure 4. Predicted secondary structures of *Baltia butleri* and *Talbotia naganum* 22 tRNA genes.

Figure 5. Predicted secondary structures of *Pontia callidice* and *Pontia daplidice* 22 tRNA genes.

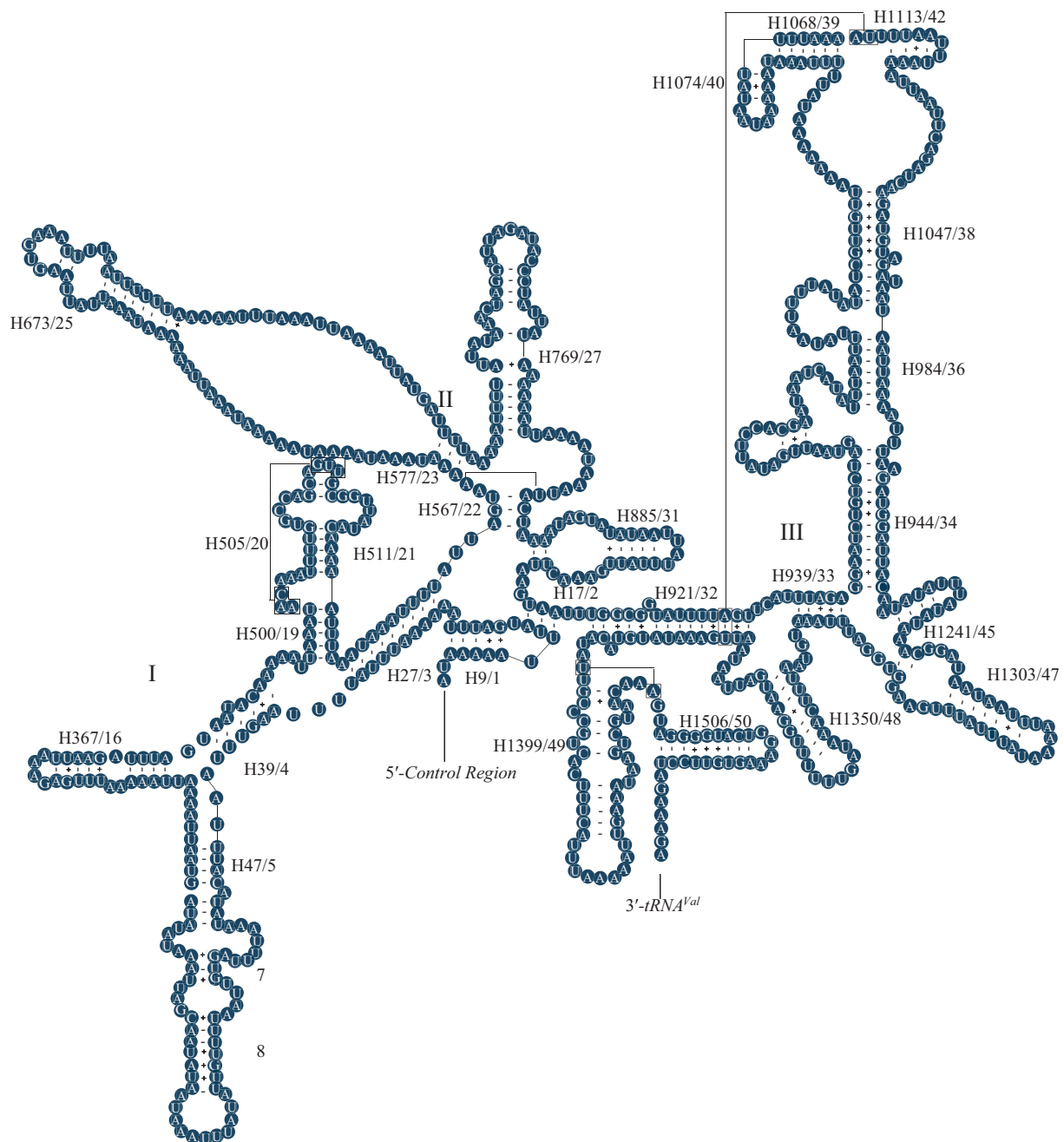
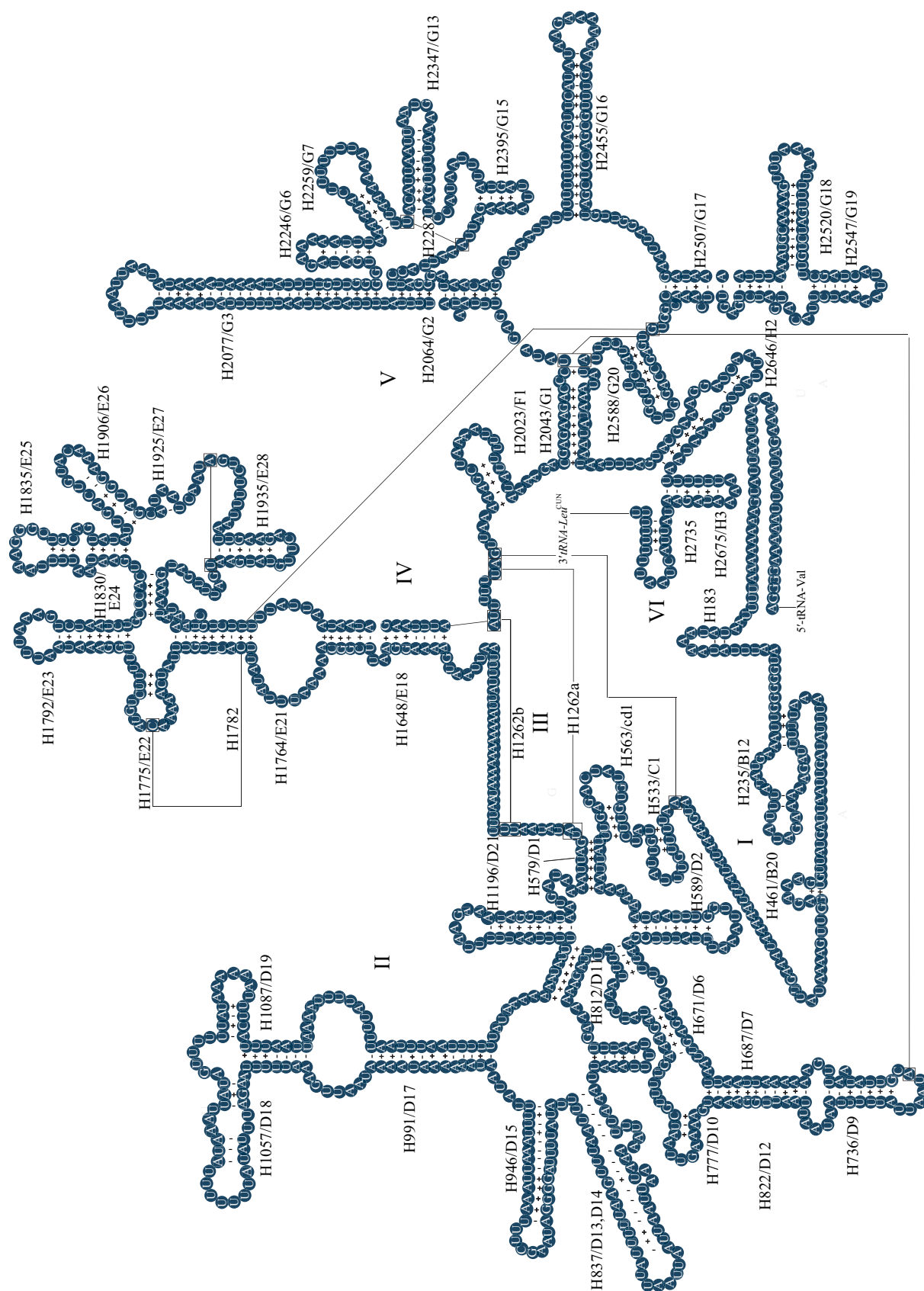


Figure 6. Predicted secondary structure of *Baltia butleri* srRNA gene.

3.5 Phylogenetic analysis

Our phylogenetic analyses showed that BI and ML phylogenetic trees generated upon the datasets of 13 PCGs, 13 PCGs + 2 rRNAs, 2 rRNAs and 2 rRNAs + 22 tRNAs all harbored the same topologies, only with slight differences of supporting values on the nodes (Fig. 15). However, the topologies of the BI and ML trees upon 22 tRNA sequence are significantly different from that upon above sequence dataset (Fig. 16). As for the deep-level phylogeny, the trees shown in Figure 15 indicated that (1) the Pieridae of this study was composed of three well-supported clades (subfamilies Dismorphiinae, Coliadinae, and Pierinae), with the relationship of (Dismorphiinae, (Coliadinae, Pierinae)), which is in congruence with a variety of previous studies based on morphological and molecular criteria (Braby *et al.*, 2006; Cao *et al.*, 2016); (2) the subfamily Dismorphiinae only represented by *Leptidea morsei* was sister to the grouping of other two subfamilies (Coliadinae and Pierinae); (3) the Coliadinae and Pierinae are monophyletic and sister to each other. However, the trees show in Figure 16 indicated that the Pieridae of this study was composed of three clades, with their relationship of

Figure 7. Predicted secondary structure of *Baltia butleri* 16S rRNA gene.

((Dismorphiinae, Pierinae), Coliadinae), which is inconsistent to a variety of previous morphological and molecular studies.

The trees shown in Figure 15 indicated that the Pierinae subfamily contained two sister tribes, the Anthocharidini and Pierini; the Anthocharidini was made up of two sister genera *Anthocharis* and *Hebomoia*; the Pierini was made up of two clades (PP = 1.00, BS = 100%): the first is the (*Pieris*, (*Baltia*, (*Pontia*, *Talbotia*))), the second is the (*Prioneris*, (*Delias*, *Aporia*)); the Coliadinae subfamily contained three genera, with their relationship of ((*Catopsilia*, *Colias*), *Gonepteryx*). However, the only tRNA based Figure 16 trees indicated that the *Leptidea* and *Anthocharis* constituted a monophyletic group which is paraphyletic to *Hebomoia*.

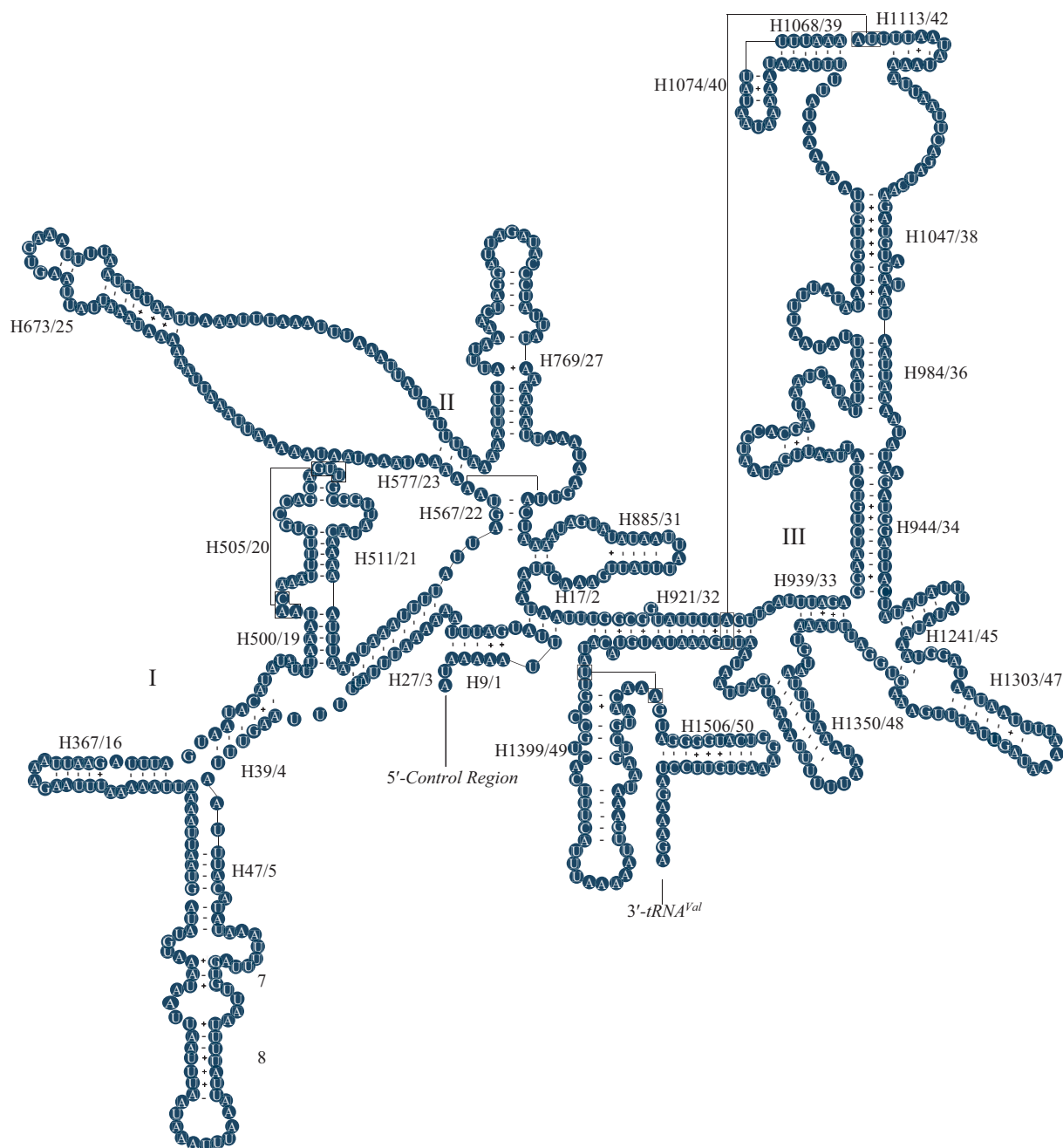
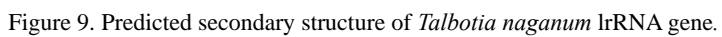


Figure 8. Predicted secondary structure of *Talbotia naganum* srRNA gene.



4 Discussion

Our phylogenetic analyses result upon mitogenomic concatenated datasets (13 PCGs, 13 PCGs+2 rRNAs) are generally consistent with those reported in a majority of previous studies (Wahlberg *et al.*, 2014; Cao *et al.*, 2016; Ding & Zhang, 2016). For instances, the phylogenetic study conducted by Wahlberg *et al.* (2014) also revealed that *Hebomoia* was sister to *Anthocharis*, and both of the two genera were clustered within the tribe Anthocharidini; the two sister genera *Pontia* and *Talbotia* was also suggested by Xu *et al.* (2007). However, some incongruences between our results and previous studies were existed herein, for example, the genus *Hebomoia* was not placed within the tribe Anthocharidini in the study conducted by Braby *et al.* (2006); our results distinctly showed that the relationship of tribe Pierini was (*Pieris* + (*Baltia* + (*Pontia* + *Talbotia*))), not in favour of those obtained by Braby *et al.* (2006) and Ding and Zhang (2016). Besides, as for the relationship among three Coliadinae genera (*Colias*, *Catopsilia*, *Gonepteryx*) of this study, our results showed that *Gonepteryx* was sister to grouping of *Colias* plus *Catopsilia*, supporting the result of Cao *et al.* (2016), but in congruence with those of Braby *et al.*

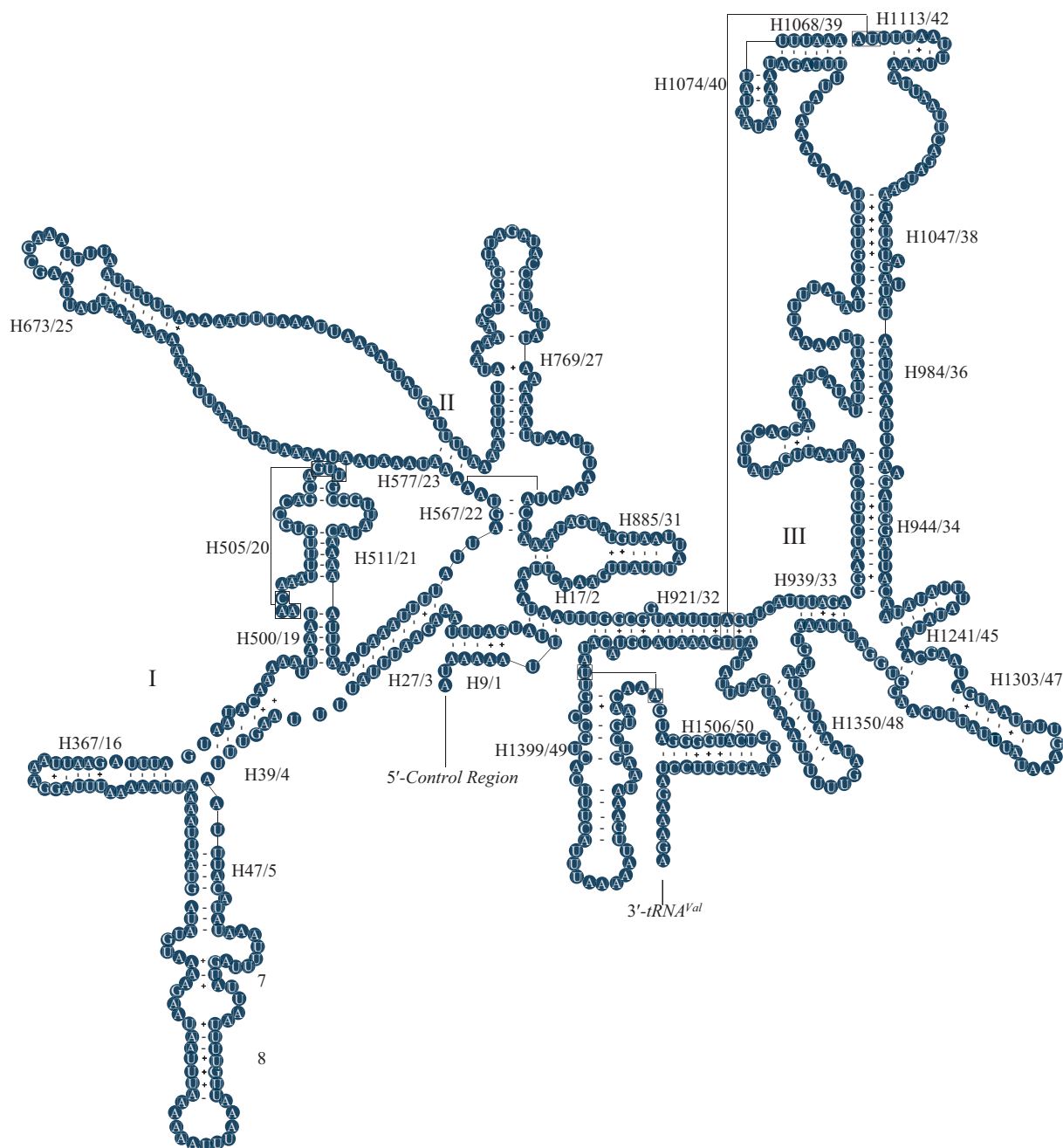
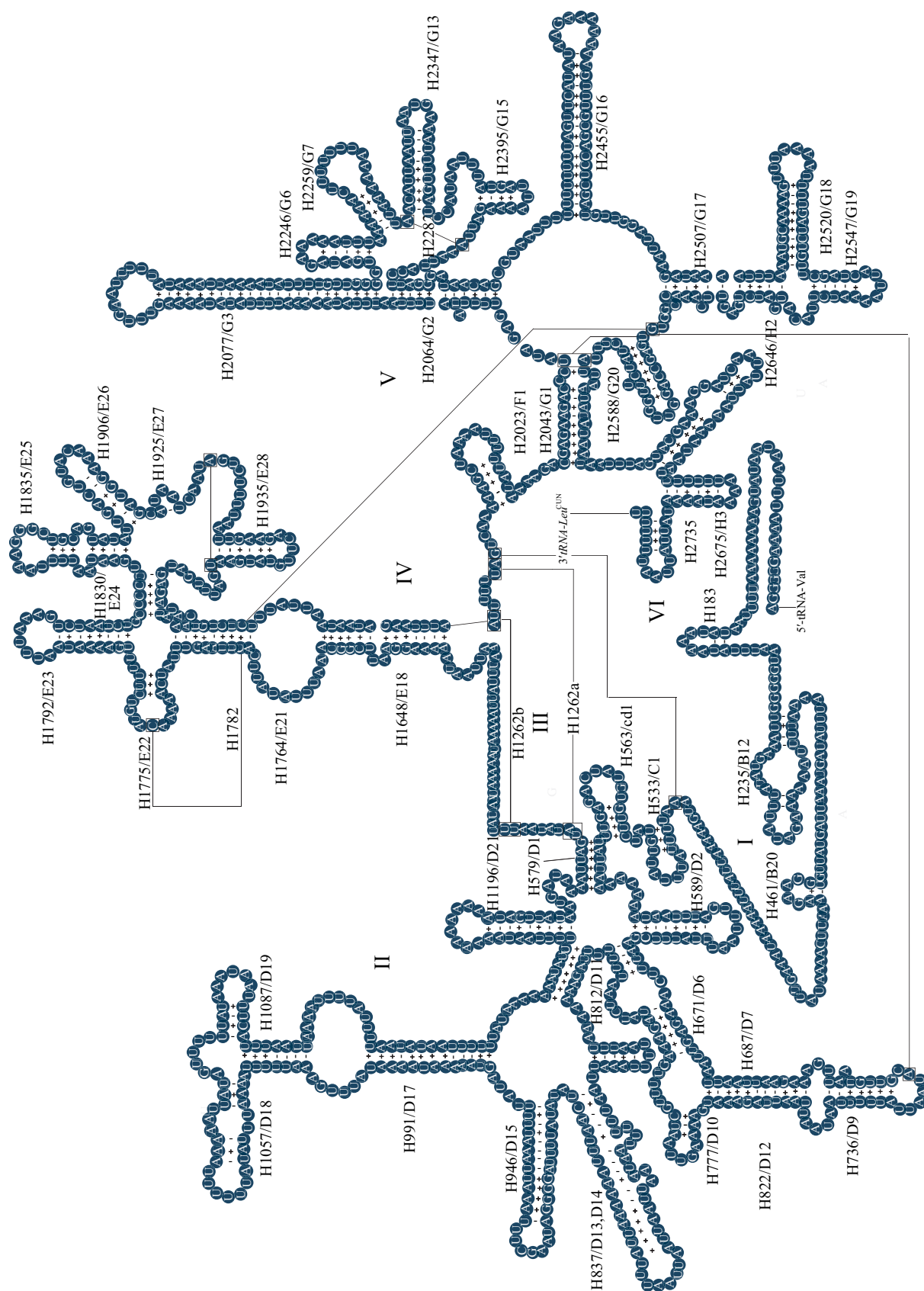


Figure 10. Predicted secondary structure of *Pontia callidice* srRNA gene.

Figure 11. Predicted secondary structure of *Pontia callidice* 16S rRNA gene.

al. (2006) and Wahlberg *et al.* (2014).

As far as we know, the relatively long sequence data are more effective in resolving the relatively higher-level phylogeny, while relatively short sequence data are often used to clarify the relatively lower-level phylogeny (Kawaguchi *et al.*, 2001; Pena *et al.*, 2006). As for the phylogenetic inference of insects and other animal groups, the target taxa at different taxonomic levels is often ready to be resolved by using different mitogenomic sequence combinations. For example, Song *et al.* (2016) reconstructed the phylogenetic trees of 199 exemplar holometabolous insect species using mitogenomic data, the results indicated that the incorporation of realigned rRNA data into protein coding sequence data could improve the resolution of relationships of deep phylogeny; Wang *et al.* (2017) conducted the mitogenome-based phylogenetic analysis of Neuropterida, and the results revealed the similar cases; Shi *et al.* (2015) reconstructed the Nymphalidae phylogeny with multiple analytic methods by using different mitogenomic dataset, the results also showed that the combination of protein coding genes plus rRNA genes are relatively powerful than other dataset in resolving the higher-level phylogenies among

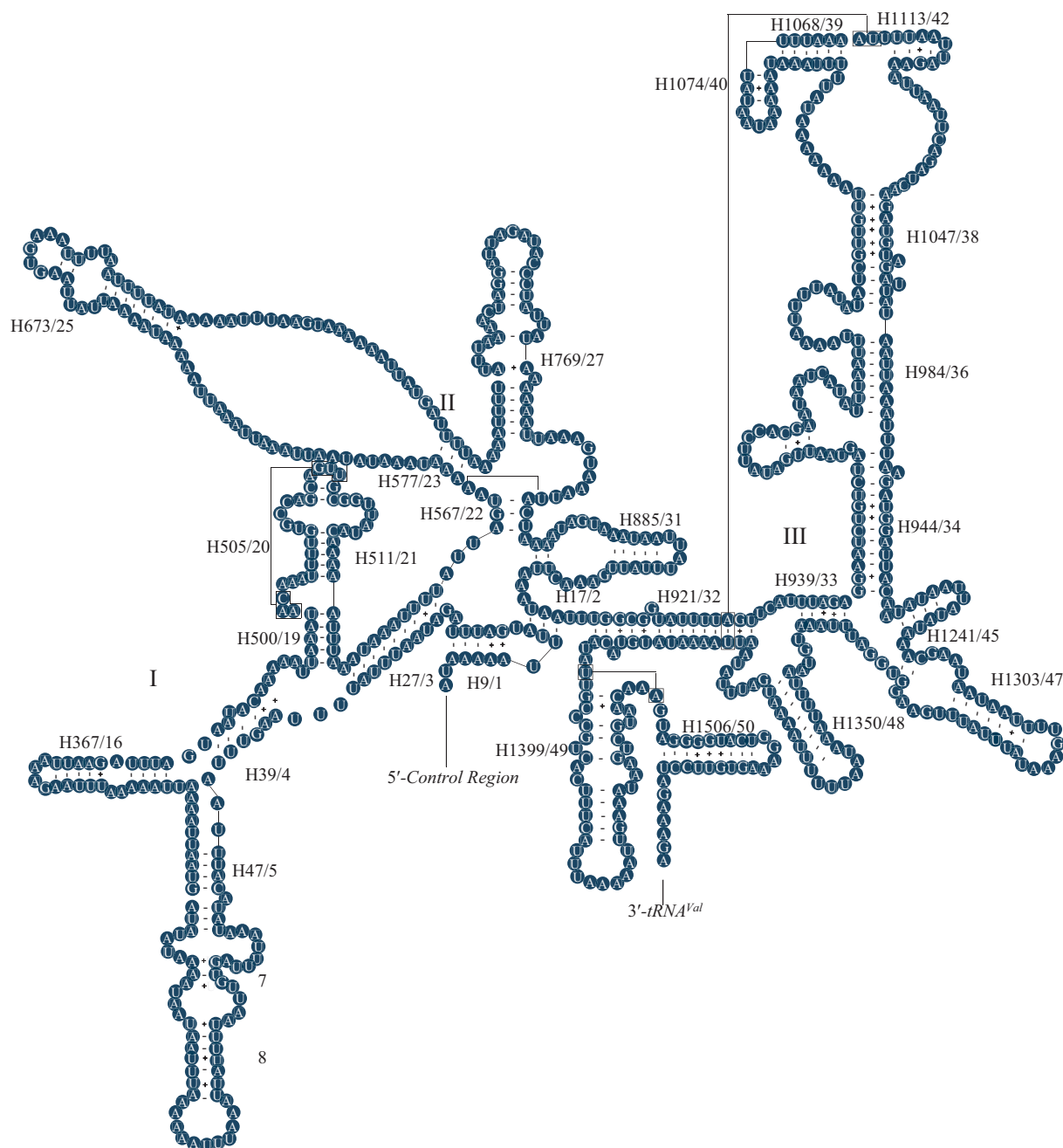
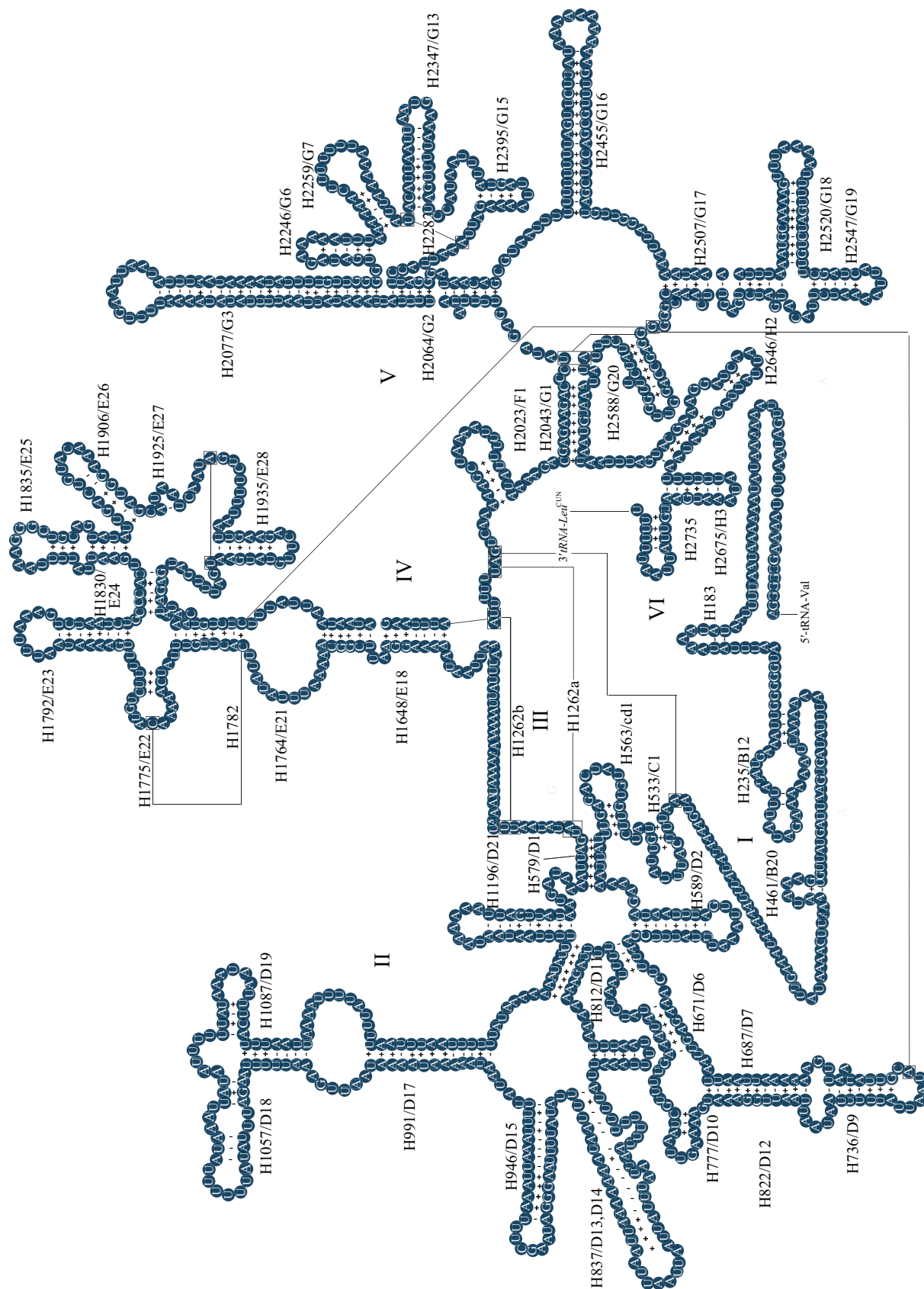


Figure 12. Predicted secondary structure of *Pontia daplidice* srRNA gene.

Figure 13. Predicted secondary structure of *Pontia daplidice* 16S rRNA gene.

lineages. In this case, the butterfly family Peridae was utilized to clarify their phylogeny with standard Bayesian inference and maximum likelihood methods by using different mitogenomic dataset, including the only tRNA sequences, the results also validated the applicability and effectiveness of protein coding gene plus rRNA gene sequences in deep phylogeny resolution. Additionally, our study showed that only tRNA sequence data were not suitable to reconstruct the deep phylogeny robustly, whereas somewhat effective in resolving lower-level phylogenies among closely related species.

Baltia butleri

srRNA -14764ATTTTAAATATTACCACATAGATTTTTTTTTTTTTTTTTATATTAAAAATTTAC
 The origin of minority strand replation(poly-T)
 ATTTATAATAAATTTTATTTAATTCTTTTATTCTTTCTCATAATATGTATATAA
 ATACAAATTCCTATTAAACAATCAATAATAAGTGTAATAAAAAAATTATAA
 ATATTATATTTATTTTTTAAACAATTTATTAAATATAAATAAATATTAATTATATA
 AAATATTAATTTATTATTAATATTTTATAATATATTAAATATTTAATATGCACGTA
 CGTATAAAAAATAAAAAGAATTTTTCTTTTATTTTTATTAAACCAATTTCAA
 TATATTTACATATAAATATAAAAAATATA15124-*tRNA*^{Met}

Talbotia naganum

srRNA -14769ATTATTTTTTTTTTTTAAATAGATTTTTTTTTTTTTTTTTATATTAAATATTTA
 The origin of minority strand replation(poly-T)
 TATTATATTAAATTTTATATAATCTTTCTTTCATTTTATAATATTTATATAAAT
 ACAAATAATTATTAAACAATAAATAAATAAATATAAATAAAAAAATAATTAA
 ATTATATATTTATATTAAAAATTAATATATATATATTTATAATTAATTATTTAA
 A micro satellite(AT)7
 ATTTTAATTTATTATTATTTTAAAAATAAATCAAATATTTAATATATATATATAT
 A micro satellite(AT)9
 ATATGTGTATACACACATACATATAATTTATTGTAATTTATTAATAAAATAAA
 CCTTATTAATTTTTTAAACCTTTTCAATAATTTATATATAAAAATAAA15155-*tRNA*^{Met}

Pontia callidice

srRNA -14748ATTTTAAATATTACCACATAGATTTTTTTTTTTTTTTTTATATTAAAAATTTAC
 The origin of minority strand replation(poly-T)
 ATTTTATTAAATTTTATTTAATTTCTTTTATTCTTTCTTAGAATATTTGTATAAAT
 ACAAATTCATTATTGAACAATTAATAATAAGTGTAATGAAAAAAATTATAAATA
 TAATATTTATTTTTTAAACAATTTATTAAATATAAATTTATAATAATTATTTAAATA
 ATTAATTTATTATTATTTTATAATATATTAAATATTTAATATGCATGCACGTATA
 AAAAAAACAAAGATAAATTTACCTTTATTTTTTTTAAACCTATTTCATAAATTT
 TACATATAAATATATAAAAAATA15109-*tRNA*^{Met}

Pontia daplidice

srRNA -14751ATTTATTAGTATTATCACATAGATTTTTTTTTTTTTTTTTATATTAAATATTTAC
 The origin of minority strand replation(poly-T)
 ATTTTTATTAAATTTTATTTAGTTCTTTTATTCTTTTCTTATAATATGTATTTA
 AATACAAATTCATTATTAAACAATTAATAATAAGTTTAAATAAAAAAATTATA
 AATATAATATTTATTTTTTAAACAATTTATTATTTTATTTAATTTATAATTATA
 TAAATATTTATATATATATATATATATTAATATATTAAATATTTAATATGCC
 A micro satellite(TA)9
 ACGCATATATATATATAAATAAATAAAGAAATTTTCTTTAATTTATTTATA
 A micro satellite(TA)6
 AACCTATTTCATAAATTTTATATATAAATATAAATAAATAAATAA15124-*tRNA*^{Met}

Figure 14. A+T-rich regions of the *Baltia butleri*, *Talbotia naganum*, *Pontia callidice* and *Pontia daplidice* mitogenomes in this study.

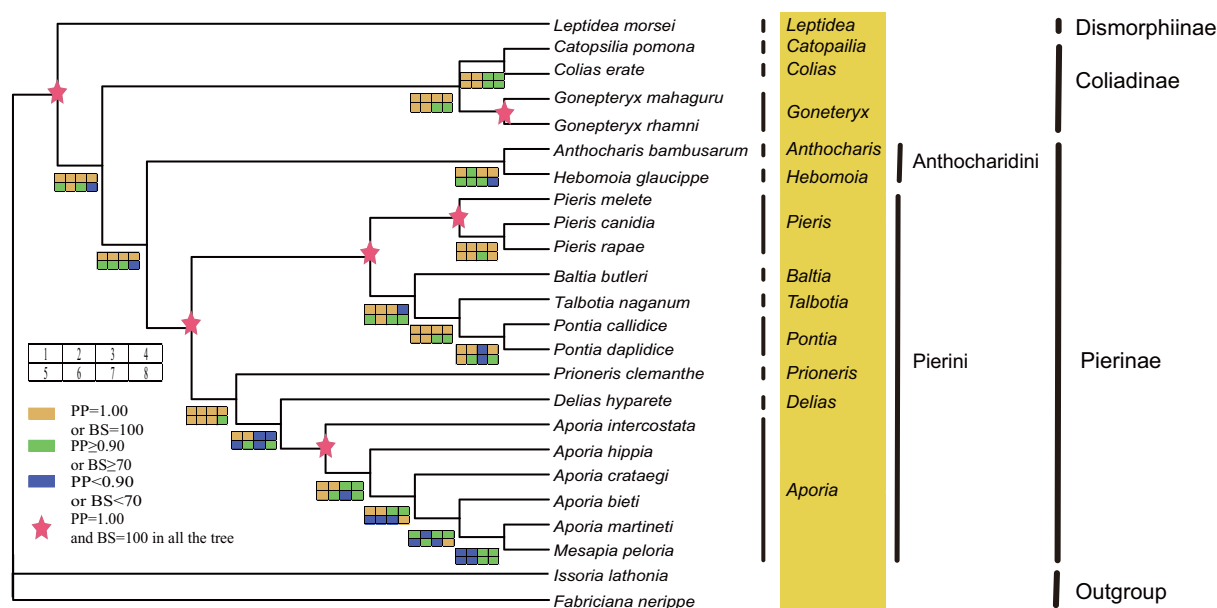


Figure 15. Bayesian inference (BI) and Maximum likelihood (ML) phylogenetic trees inferred from mitochondrial genomes of pierid family based on four datasets (13PCGs, 13PCGs+2rRNAs, 2rRNAs, 2rRNAs+22tRNAs).

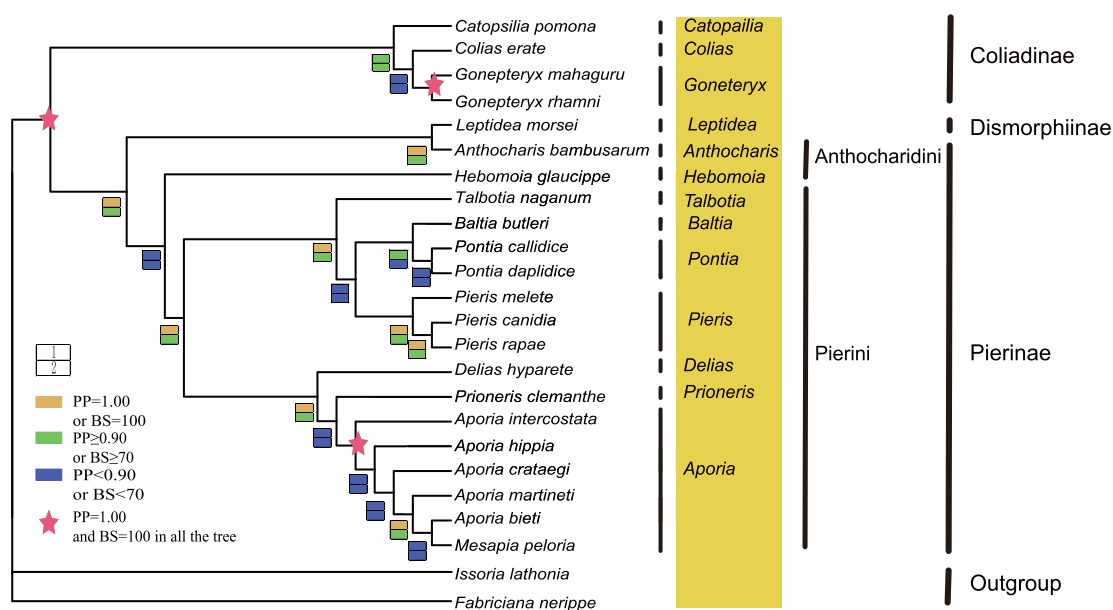


Figure 16. Bayesian inference (BI) and Maximum likelihood (ML) phylogenetic trees inferred from mitochondrial genomes of pierid family based on 22tRNA genes.

Fundings This work was supported by the National Natural Science Foundation of China (41172004, 41472028), and the funds from the State Key Laboratory of Palaeobiology and Stratigraphy (Nanjing Institute of Geology and Palaeontology, CAS) (Y626040108).

References

- Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, 27(2): 573–580.
 Boore, J.L. 1999. Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8): 1767–1780.

- Braby, M.F., Trueman, J.W.H. 2006. Evolution of larval host plant associations and adaptive radiation in pierid butterflies. *Journal of Evolutionary Biology*, 19(5): 1677–1690.
- Braby, M.F., Vila, R., Pierce, N.E. 2006. Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zoological Journal of the Linnean Society*, 147(2): 239–275.
- Brown, W.M., George, M.J., Wilson, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76(4): 1967–1971.
- Cameron, S.L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology*, 59(1): 95–117.
- Cameron, S.L., Whiting, M.F. 2008. The complete mitochondrial genome of the Tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae): and an examination of mitochondrial gene variability within butterflies and moths. *Gene*, 408(2): 112–123.
- Cao, Y., Hao, J.S., Sun, X.Y., Zheng, B., Yang, Q. 2016. Molecular phylogenetic and dating analysis of pierid butterfly species using complete mitochondrial genomes. *Genetics and Molecular Research*, 15(4): gmr15049196.
- Clench, H.K. 1955. Revised classification of the butterfly family Lycaenidae and its allies. *Annals of the Carnegie Museum*, 33: 261–274.
- Ding, C., Zhang, Y. 2016. Phylogenetic relationships of Pieridae (Lepidoptera: Papilionoidea) in China based on seven gene fragments. *Entomological Science*, 20(1): 15–23.
- Dowton, M., Castro, L.R., Austin, A D. 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. *Invertebrate Systematics*, 16(3): 345–356.
- Ehrlich, P.R. 1958. The comparative morphology, phylogeny and higher classification of the butterflies (Lepidoptera: Papilionoidea). *Kansas University Science Bulletin*, 39: 305–370.
- Ehrlich, P.R., Ehrlich, A.H. 1967. The phenetic relationships of the butterflies I. Adult taxonomy and the nonspecificity hypothesis. *Systematic Zoology*, 16(4): 301–317.
- Fang, J., Wu, Y., Wang, H., Sun, Z., Han, D., Zhang, B. 2015. The complete nucleotide sequence of the mitochondrial genome of the Oriental Pieris, *Pieris canidia* (Lepidoptera: Pieridae). *DNA Sequence*, 27(6): 4088–4089.
- Hall, T.A. 1999. BioEdit : a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41(41): 95–98.
- Hao, J.J., Hao, J.S., Sun, X.Y., Zhang, L.L., Yang, Q. 2014. The complete mitochondrial genomes of the Fenton's wood white, *Leptidea morsei*, and the lemon emigrant, *Catopsilia pomona*. *Journal of Insect Science*, 14(130). doi: 10.1093/jis/14.1.130.
- Hao, J.J., Wang, Y.L., Sun, X.Y., Zhang, L.L., Hao, J.S., Yang, Q. 2013. The complete mitochondrial genome of *Hebomoia glaucippe* (Lepidoptera: Pieridae). *Mitochondrial DNA*, 24(6): 668–670.
- Hao, J.S., Li, C.X., Sun, X.Y., Yang, Q. 2005. Phylogeny and divergence time estimation of cheilostome bryozoans based on mitochondrial 16S rRNA sequences. *Chinese Science Bulletin*, 50(12): 1205–1211.
- Jiang, S.T., Hong, G.Y., Yu, M., Li, N., Yang, Y., Liu, Y.Q., Wei, Z.J. 2009. Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae). *International Journal of Biological Sciences*, 5(4): 351–365.
- Kawaguchi, A., Miya, M., Nishida, M. 2001. Complete mitochondrial DNA sequence of *Aulopus japonicus* (Teleostei: Aulopiformes): a basal eurypterygii: longer DNA sequences and higher-level relationships. *Ichthyological Research*, 48(3): 213–223.
- Kim, M.I., Baek, J.Y., Kim, M.J., Jeong, H.C., Kim, K.G., Bae, C.H., Kim, I. 2009. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other Lepidopteran insects. *Molecules and Cells*, 28(4): 347–363.
- Kumar, S., Stecher, G., Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7): 1870–1874.
- Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6): 1695–1701.
- Lavrov, D.V., Brown, W.M., Boore, J.L. 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proceedings of the National Academy of Sciences of the United States of America*, 97(25): 13738–13742.
- Li, H., Shao, R.F., Song, N., Song, F., Jiang, P., Li, Z.H., Cai, W.Z. 2015. Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Scientific Reports*, 5(8527). doi: 10.1038/srep08527.
- Liu, Q.N., Zhu, B.J., Dai, L.S., Liu, C.L. 2013. The complete mitogenome of *Bombyx mori* strain Dazao (Lepidoptera: Bombycidae) and comparison with other Lepidopteran insects. *Genomics*, 101(1): 64–73.
- Lowe, T.M., Eddy, S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 25(5): 955–964.
- Mao, Z.H., Hao, J.S., Zhu, G.P., Hu, J., Si, M.M., Zhu, C.D. 2010. Sequencing and analysis of the complete mitochondrial genome of *Pieris rapae* linnaeus (Lepidoptera: Pieridae). *Acta Entomologica Sinica*, 53(11): 1295–1304.
- Park, J.S., Cho, Y., Min, J.K., Nam, S.H., Kim, I. 2012. Description of complete mitochondrial genome of the black-veined white, *Aporia crataegi*, (Lepidoptera: Papilionoidea), and comparison to papilionoid species. *Journal of Asia-Pacific Entomology*, 15(3): 331–341.

- Pena, C., Wahlberg, N., Weingartner, E., Kodandaramaiah, U., Nylin, S., Freitas, A.V.L., Brower, A. 2006. Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. *Molecular Phylogenetics and Evolution*, 40(1): 29–49.
- Pollock, D.D., Watt, W.B., Rashbrook, V.K., Iyengar, E.V. 1998. Molecular phylogeny for *Colias* butterflies and their relatives (Lepidoptera: Pieridae). *Annals of the Entomological Society of America*, 91(5): 524–531.
- Ronquist, F., Teslenko, M., Van, d.M.P., Ayres, D.L., Darling, A., Höhna, S., Huelsenbeck, J.P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3): 539–542.
- Shadel, G.S., Clayton, D.A. 1993. Mitochondrial transcription initiation. variation and conservation. *Journal of Biological Chemistry*, 268(22): 16083–16086.
- Shi, Q.H., Sun, X.Y., Wang, Y.L., Hao, J.S., Yang, Q. 2015. Morphological characters are compatible with mitogenomic data in resolving the phylogeny of Nymphalid butterflies (Lepidoptera: Papilionoidea: Nymphalidae). *PLoSOne*, 10(4): e0124349.
- Shi, Q.H., Xia, J., Sun, X.Y., Hao, J.S., Yang, Q. 2012. Complete mitogenome of the *Painted jezebel*, *Delias hyparete* Linnaeus (Lepidoptera: Pieridae) and its comparison with other butterfly species. *Zoological Systematics*, 33(e5): 111–120.
- Silvestro, D., Michalak, I. 2012. raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity and Evolution*, 12(4): 335–337.
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, 87(6): 651–701.
- Singh, V.K., Mangalam, A.K., Dwivedi, S., Naik, S. 1998. Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques*, 24(2): 318–319.
- Song, F., Li, H., Jiang, P., Zhou, X.G., Liu, J.P., Sun, C.H., Vogler, A.P., Cai, W.Z. 2016. Capturing the phylogeny of holometabola with mitochondrial genome data and bayesian site-heterogeneous mixture models. *Genome Biology and Evolution*, 8(5): 1411–1426.
- Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21): 2688–2690.
- Su, C.Y., Shi, Q.H., Sun, X.Y., Ma, J.Y., Li, C.X., Hao, J.S., Yang, Q. 2017. Dated phylogeny and dispersal history of the butterfly subfamily Nymphalinae (Lepidoptera: Nymphalidae). *Scientific Reports*, 7(8799). doi.org/10.1038/s41598-017-08993-w.
- Sun, Q.Q., Sun, X.Y., Wang, X.C., Gai, Y.H., Hu, J., Zhu, C.D., Hao, J.S. 2012. Complete sequence of the mitochondrial genome of the Japanese buff-tip moth, *Phalera flavescentis* (Lepidoptera: Notodontidae). *Genetics and Molecular Research*, 11(4): 4213–4225.
- Sun, X.Y., Shao, L.L., Peng, C.M., Hao, J.S., Yang, Q. 2015. The complete mitochondrial genome of *Eurema hecabe* (Lepidoptera: Pieridae: Coliadinae). *Mitochondrial DNA*, 26(5): 783.
- Sun, Y.X., Wang, L., Wei, G.Q., Qian, C., Dai, L.S., Sun, Y., Abbas, M.N., Zhu, B.J., Liu, C.L. 2016. Characterization of the complete mitochondrial genome of *Leucoma salicis* (Lepidoptera: Lymantriidae) and comparison with other lepidopteran insects. *Scientific Reports*, 6(39153). doi:10.1038/srep39153.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25(25): 4876–4882.
- Wahlberg, N., Rota, J., Braby, M.F., Pierce, N.E., Wheat, C.W. 2014. Revised systematics and higher classification of pierid butterflies (Lepidoptera: Pieridae) based on molecular data. *Zoologica Scripta*, 43(6): 641–650.
- Wang, Y.Y., Liu, X.Y., Garzón-Orduña, I.J., Winterton, S.L., Yan, Y., Aspöck, U., Aspöck, H., Yang, D. 2017. Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics*, 33(6): 617–636.
- Wu, Y.H., Fang, J., Li, W.B., Han, D.M., Wang, H., Zhang, B.W. 2015. The complete mitochondrial genome of *Colias erate* (Lepidoptera: Pieridae). *Dna Sequence*, 27(6): 4209–4210.
- Xu, L., Hao, J.S., Zhu, G.P., Yin, X.B., Pan, H.C., Huang, D.Y., Zhang, X.P. 2007. Molecular phylogenetic relationships of some species and genera in Pierinae and Coliadinae (Pieridae) based on partial sequence of mitochondrial COI and Cytb genes. *Acta Zootaxonomica Sinica*, 32(1): 842–850.
- Yang, J.N., Xu, C., Li, J.L., Lei, Y., Fan, C., Gao, Y., Xu, C.R., Wang, R.J. 2014. The complete mitochondrial genome of *Gonepteryx mahaguru* (Lepidoptera: Pieridae). *Mitochondrial Dna Part A, DNA Mapping Sequencing and Analysis*, 27(2): 877–878.
- Yang, L., Wei, Z.J., Hong, G.Y., Jiang, S.T., Wen, L.P. 2009. The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae). *Molecular Biology Reports*, 36(6): 1441–1449.
- Zhang, L.L., Hao, J.S., Huang, D.Y., Sun, X.Y., Hao, J.J., Peng, C.M., Yang, Q. 2013. Complete mitochondrial genomes of the bright sunbeam *Curetis bulis* and the small copper *Lycaena phlaeas* (Lepidoptera: Lycaenidae) and their phylogenetic implications. *Genetics and Molecular Research*, 12(4): 4434–4445.